ARCHIVES NATIONALES D'OUTRE-MER

AMELIORATION DU PORTAIL IREL

Cahier des clauses techniques particulières

Sommaire

Présentation générale....3

Situation actuelle....5

Description générale des travaux demandés....9

Nouvelles fonctionnalités à intégrer dans IREL....10

Amélioration de la recherche transversale par la base géographique aux inventaires XML/EAD....12

Amélioration de la recherche transversale par vocabulaire aux inventaires XML/EAD....15

Modification de l'outil attribuant les identifiants ARK....16

Intégration d'un outil de gestion et de publication des notices d'autorités encodées en XML/EAC-CPF....17

Amélioration de la recherche transversale par la base nominative « Personnes et familles »..19

Amélioration de l'application de l'Etat civil....20

Présentation générale

Les Archives nationales d'outre-mer

Situées à Aix-en-Provence depuis 1966, les Archives nationales d'outre-mer conservent les archives publiques de la présence coloniale française outre-mer. S'y ajoutent des archives privées et d'entreprises relatives à l'outre-mer ainsi qu'une bibliothèque, une cartothèque et une iconothèque spécialisées.
Elles ont notamment pour mission, outre leur conservation, la communication de ces archives aux chercheurs quelle que soit leur qualité (universitaire, historien, généalogiste, etc.). La consultation de ces documents se fait sur place, dans les salles de lecture spécialement aménagées.
La totalité des documents conservés par les Archives nationales d'outre-mer occupe actuellement 38 kilomètres de rayonnages.

Les instruments de recherche

Pour pouvoir consulter un document, encore faut-il savoir qu'il existe. Pour en signaler l'existence au chercheur, les archives en élaborent des descriptions qui sont recensées dans des "instruments de recherche". Ceux-ci se présentent matériellement sous diverses formes : livres, classeurs, fichiers (tiroirs) et depuis peu fichiers informatiques. A l'exception de ces derniers ils peuvent être imprimés, dactylographiés ou manuscrits. Ils se distinguent également par leur niveau de description selon une terminologie "métier" spécifique, mais qui différencie nettement un niveau de description général qui donne des informations globales (l'état général des fonds) et un niveau de description plus ou moins fin, pouvant aller jusqu'à la pièce, au niveau de chaque carton ou registre (les inventaires).
Ces instruments de recherche sont donc indispensables au chercheur pour accéder aux archives. En effet, seule leur consultation permet de déterminer dans quels services des documents susceptibles d'intéresser une recherche précise sont conservés, s'ils sont librement accessibles, et comment en obtenir la communication.

La politique de normalisation des instruments de recherche

L'élaboration de ces instruments de recherche doit répondre, depuis 1994, à une norme professionnelle internationale gérée par le Conseil international des Archives : la norme

ISAD(G) qui préconise la description à plusieurs niveaux. Pour en faciliter la diffusion et l'application, une DTD a été mise au point pour tous les instruments de recherche archivistiques encodés en XML : l'EAD.
De même, le Conseil international des Archives a promu une autre norme pour décrire les producteurs d'archives : la norme ISAAR(CPF). Celle-ci a, à son tour, donné naissance à une nouvelle DTD devenue depuis un schéma pour les notices d'autorités réalisées en conformité avec elle et encodées en XML : l'EAC.
Les Archives nationales d'outre-mer appliquent réglementairement ces normes et constituent des inventaires et des notices d'autorité encodés en XML/EAD et XML/EAC.

Le portail d'accès aux instruments de recherche en ligne (IREL)

Le lectorat qui fréquente les salles de lecture des Archives nationales d'outre-mer est très majoritairement d'origine extérieure à la région Provence-Alpes-Côte-d'Azur et compte de nombreux étrangers venant de tous les continents. Il souhaite donc tout naturellement bénéficier du maximum d'informations sur les archives conservées à Aix-en-Provence le plus en amont possible, avant même son déplacement, pour mieux l'organiser et le rentabiliser.
C'est pour satisfaire cette demande légitime que les Archives nationales d'outre-mer ont mis en place sur leur site Internet en 2004 un portail d'accès aux instruments de recherche en ligne (IREL) décrit ci-après.
A cette époque les informations de niveau général (état général des fonds) ont fait l'objet d'une première publication. Ce portail a par la suite été enrichi par la mise en ligne de bases de données (banque d'images Ulysse, état civil, catalogue de la bibliothèque) et, depuis son évolution intervenue en 2008, d'instruments de recherche encodés en EAD.

Situation actuelle

Pour mieux comprendre l'état actuel du portail IREL, voici donc une présentation du projet depuis 2008 et son architecture technique.

L'objet du marché de 2008 (IREL 2) était d'accompagner techniquement la recomposition de la politique documentaire de l'institution. Avec l'expérience, les bases de données conservent toute leur pertinence pour certaines informations très tabulaires, comme des bases nominatives.

Le tout-XML n'est plus conçu comme un objectif, il en résulte le besoin de coordonner des sources multiples d'informations. Par ailleurs, le développement des instruments de recherche EAD devient moins monolithique. L'informatique tend à préférer le modèle de l'arbre unique qui se déploie dans le moindre détail des fonds. Cet ordre conduit à vouloir unifier un seul principe d'organisation. La pratique de l'archiviste, ainsi que les besoins du chercheur, amènent les ANOM à considérer que les principes d'organisation sont nécessairement pluriels : les institutions, mais aussi les aires géographiques, les périodes historiques, ou les personnes physiques et morales. Il en résulte plusieurs axes de développement des inventaires, multipliant le nombre de racines indépendantes, et les croisements. IREL 2 est la mise en technique de cette réflexion, changeant en profondeur les principes de développement d'IREL 1.

Architecture technique

Puisque les sources de données seront désormais multiples : XML (EAD, EAC), bases de données (PHP/MySQL), le portail n'est plus un *Système Documentaire XML* (SDX), mais un développement spécifique et optimisé pour assurer une réponse performante au public. IREL est conçu comme un portail d'agrégation de sources d'informations multiples, il est écrit en PHP, employant XML ou SQL au mieux de leurs performances.
Cette architecture en deux couches permet de s'accommoder d'un problème de stabilité constaté sur les applications conçues en framework. L'accumulation des couches d'abstraction a un coût en performances, et augmente les risques d'effets de bord imprévisibles. Le portail IREL en PHP est au moins assuré de toujours répondre au public.
En 2010, ont été développés un indexeur autonome produisant les index Lucene à partir des composants EAD et une application de recherches utilisant les index Lucene.

La recherche dans les inventaires détaillés encodés en XML/EAD

Comme décrit précédemment, IREL est un portail regroupant un ensemble d'applications autonomes, chacune étant dotée de ses propres moyens de recherche :

- Inventaires en XML/EAD
- Bibliothèque
- Banque d'images ULYSSE
- Etat-civil

Il est cependant nécessaire de détailler les outils développés pour la recherche dans les inventaires détaillés encodés en XML/EAD pour appréhender certaines modifications demandées.

Recherche sur un seul inventaire

Le choix d'un inventaire particulier s'opère à partir de la page d'accueil « Liste des inventaires ». L'organisation de la présentation des inventaires accessibles (par territoire, par producteurs, par période chronologique, par onglets spécifiques) relève des ANOM qui disposent des outils nécessaires inclus dans le programme de gestion d'IREL pour les inventaires en XML/EAD.
La consultation de l'inventaire se fait en « dépliant » les branches de l'arbre XML. La recherche s'effectue au moyen, au choix du lecteur, d'un formulaire de recherche « simple » (plein texte, lieux, dates) ou d'un formulaire de recherche « avancé » qui ajoute au formulaire « simple » tous les champs d'indexation (outres lieux et dates) intégrés dans le fichier XML.

Recherche transversale à tous les inventaires

Le chercheur disposera à terme de 4 moyens d'accès aux informations contenues dans les inventaires en fonction de ses centres d'intérêt. Certains sont déjà opérationnels, d'autres sont en cours de constitution :

- Formulaires de recherche (opérationnels)
- Base nominative « Personnes et familles » (opérationnelle)
- Vocabulaire (données en cours de constitution)
- Base géographique (données en cours de constitution)

<u>Formulaires de recherche</u>

Comme pour la recherche sur un seul inventaire, deux formulaires de recherche sont disponibles dès la page d'accueil : formulaire « simple » identique, et formulaire « avancé » comportant actuellement les champs de recherche suivants en plus de ceux du formulaire « simple » :

- Institutions
- Ordres religieux
- Compagnies commerciales
- Parti/Mouvement politique
- Association
- Peuple/Ethnie

- Région ou pays d'origine/Nationalité
- Bateaux
- Associations

Ces champs de recherche peuvent être associés à un nom de lieu ou à une date. Les résultats sont présentés dans une « fenêtre des résultats », dans l'ordre des inventaires répondant par au moins une occurrence à(aux) la question(s) posée(s). Le nombre d'occurrences concernées est précisé pour chaque inventaire. Le chercheur accède alors au premier inventaire qu'il souhaite consulter, puis revient à la « fenêtre des résultats » pour consulter le suivant, etc.

Base nominative « Personnes et familles »

Cette base permet l'accès à toutes les personnes citées dans les inventaires, en les identifiant le plus précisément possible, à partir des patronymes, complétés éventuellement de prénom(s). Le résultat renvoie toutes les occurrences concernant une personne identifiée dans tous les inventaires où elle apparaît, sans risque d'être confondue avec un homonyme, puisque à chaque personne dument identifiée est attribuée un identifiant informatique.
Il convient de noter que cette base ne renvoie pas uniquement les résultats des fichiers encodés en XML/EAD, mais également ceux des bases de données complémentaires aux instruments de recherche en XML/EAD, comme la base nominative des condamnés aux bagnes coloniaux qui complète l'inventaire du bagne.

Vocabulaire

Cet outil de recherche n'est pas encore accessible au public. L'élaboration du vocabulaire contrôlé (dictionnaire et domaines) est en cours pour l'Ancien Régime. Dès ce premier stade achevé, il sera mis à disposition, le vocabulaire des XIXe et XXe siècle devant suivre par la suite.
Les Archives nationales d'outre-mer ont souhaité offrir au chercheur une recherche guidée dans le vocabulaire des instruments de recherche déclinée de la façon suivante :

- Un accès à un terme retenu par un dictionnaire
- Un accès à un domaine regroupant les termes relevant de ce domaine.

Le choix des termes à indexer relève de l'archiviste responsable du vocabulaire utilisé dans les instruments de recherche. Ils sont retenus en fonction de leur intérêt pour le chercheur selon l'inventaire encodé en EAD.
Chaque terme est intégré dans un dictionnaire structuré de la façon suivante : vedette, formes anciennes éventuelles, définition, domaine(s) de rattachement (avec possibilité de définitions spécifiques expliquant les différentes acceptions lors de relations avec plusieurs domaines).
Chaque terme indexé, qui peut éventuellement être une expression, est rattaché à au moins un domaine.
La liste des domaines est définie par l'administrateur du vocabulaire. Elle n'est pas figée à un instant "T" et est susceptible d'évolution. Tous les termes d'indexation relevant d'un domaine y sont associés. On peut donc trouver le même terme dans plusieurs domaines. Il n'y a pas de sous-domaines.

Exemple :

café, cafés

- Dom. **Végétaux**

Graine du caféier. Par extension, employé pour caféier.

- Dom. **Agriculture**

Culture répandue dans les colonies d'Amérique, d'Afrique et d'Asie.

- Dom. **Commerce**

Produit exporté vers l'Europe des colonies d'Amérique, d'Afrique et d'Asie à partir du XVIIe siècle et objet d'échanges commerciaux importants.

- Dom. **Fiscalité**

Des droits de douane pesaient sur l'exportation et l'importation du café.

- Dom. **Alimentation, approvisionnements**

Boisson aux propriétés stimulantes et toniques obtenue par l'infusion des graines du caféier torréfiées et moulues.

Le chercheur peut choisir de sélectionner le terme associé à son/ses domaine(s) et les placer dans un panier. Dans ce dernier cas, la recherche ne porte que sur les occurrences répondant à ce domaine ou ces domaines.

Base géographique

Cet outil de recherche n'est pas encore accessible au public. L'élaboration de la base de connaissance est en cours et devrait être mise en service prochainement.

Les Archives nationales d'outre-mer ont découplé l'accès aux informations géographiques contenues dans les instruments de recherche en séparant :

1) l'indexation géographique (comment trouver toutes les occurrences d'un même lieu dénommé de la même façon dans les instruments de recherche) qui sert à la recherche via le champ « Lieu » des formulaires de recherche

2) un accès à partir de critères géographiques divers : points plus ou moins définis (fort, ville, etc.), aires naturelles (baie, île, etc.) ou artificielles (pays, territoire, circonscription administrative, etc.), lignes (route, chemin de fer, etc.).

Pour assurer l'homogénéité et la cohérence de l'indexation effectuée, un outil d'assistance a été développé par les Archives nationales d'outre-mer.

A partir de cette indexation, a été développé un outil permettant au chercheur la possibilité d'effectuer une recherche à partir d'un nom géographique en le restituant dans ses différents contextes et en fournissant les explications historiques lui permettant à la fois de mieux définir sa requête et de comprendre les résultats obtenus.

Autres fonctionnalités rajoutées

- Possibilité d'ajouter une préface aux instruments de recherches.
- Mise en ligne de la version PDF des instruments de recherche.
- Surlignage de(s) terme(s) recherchés dans l'affichage des réponses.
- Intégration des identifiants ARK aux fichiers XML-EAD par traitement automatique.

Description générale des travaux demandés

Présentation de l'opération

Les Archives nationales d'outre-mer souhaitent améliorer leur portail d'accès aux instruments de recherche en ligne IREL afin d'une part d'en faciliter la compréhension et la prise en main par l'internaute, d'autre part de perfectionner les outils de recherche transversaux aux inventaires encodés en XML/EAD.

Travaux demandés

Les travaux demandés sont présentés dans l'ordre de priorité qui correspond aux attentes des Archives nationales d'outre-mer.

1) Nouvelles fonctionnalités à intégrer dans IREL
2) Amélioration de la recherche transversale par la base géographique aux inventaires XML/EAD
3) Amélioration de la recherche transversale par vocabulaire aux inventaires XML/EAD
4) Modification de l'outil attribuant les identifiants ARK
5) Intégration d'un outil de gestion et de publication des notices d'autorités encodées en XML/EAC-CPF dans IREL.
6) Amélioration de la recherche transversale par la base nominative « Personnes et familles »
7) Amélioration de l'application de l'état civil

Nouvelles fonctionnalités à intégrer dans IREL

1) Proposer l'accès à un mode d'emploi général d'IREL (recherche transversale, formulaires général et particuliers, accès spécifiques) qui devra pouvoir être rédigé et modifié par les ANOM.
2) Proposer un avertissement permettant de connaître l'état d'avancement des différentes bases (noms de personnes, noms de lieux , vocabulaire). Il s'agit, dès la page d'accueil de ces bases, de signaler au chercheur les instruments de recherche publiés dont les éléments n'ont pas encore été intégrés dans ces bases. Cet avertissement doit pouvoir être rédigé et modifié par les ANOM au fur et à mesure de l'avancement des travaux de publication et d'indexation.
3) Possibilité de mettre en ligne un complément en version PDF (chronologie, avant-propos, tableau, etc.) relié à un fichier XML-EAD. Comme pour la préface, il serait souhaitable de signaler au chercheur l'existence d'autres documents élaborés pour compléter un instrument de recherche et qui n'ont pas leur place dans l'instance EAD. Ils pourraient du reste apparaître et être téléchargés de façon identique à ce qui a été fait pour une préface.
4) Page d'accès au portail : évolution de l'outil permettant la modification de cette page en proposant un graphisme et une ergonomie nouvelle. Les informations de la page d'accès au portail sont rédigées et modifiées par les ANOM. Sans revenir sur cette fonctionnalité, il serait souhaitable d'en revoir l'ergonomie et la présentation graphique.
5) Modification de l'outil permettant l'organisation de la page d'accès aux inventaires encodés en XML/EAD par les ANOM. La page d'accueil de l'accès aux inventaires encodés en XML/EAD propose des listes de fonds organisés selon différents classements hiérarchiques. Ces listes sont personnalisables dans l'interface d'administration de l'application IREL. Or l'accroissement des inventaires mis en ligne rendant leur visibilité moindre, il conviendrait de pouvoir obtenir une ergonomie plus satisfaisante et une arborescence dynamique. Il serait également souhaitable de disposer :
    a. D'un nuage de mots généré à partir des entités nommées en page d'accueil.
    b. D'un avertissement et d'informations sur l'usage de cette partie d'IREL sous forme de bouton.
6) Modifier l'affichage de la présentation des éléments d'information encodés dans la balise « list » de l'EAD lors de la consultation des inventaires encodés en XML/EAD. En effet, ils apparaissent alignés trop à gauche, ce qui provoque parfois une difficulté de compréhension de l'inféodation de ces informations à une information supérieure, notamment dans le cas d'un élément « list » inclus dans un élément « p ». Il serait

souhaitable, par un alignement décalé à droite, de visualiser cette hiérarchie.

7) Rendre plus visible le bouton « Aide » dans l'ensemble du portail IREL.
8) Page d'accueil d'un IR : la rubrique suit l'ISAD-G. Pour la plupart des inventaires, les auteurs rédigent des textes d'information sous des rubriques intitulées par exemple « Introduction », « Intérêt historique », « Sources complémentaires », « Histoire administrative » qui correspondent à des balises « bioghist », « custodist », « scopecontent », « relatedmaterial », « separatedmaterial ». Certains textes très riches sont également très longs et actuellement leur lecture se fait en parcourant la page grâce à un ascenseur. Il faudrait ne faire apparaître que 2 à 300 signes (soit l'équivalent d'un court paragraphe ou d'un début de paragraphe) et proposer au lecteur la lecture de l'ensemble du texte concerné par un bouton « Pour voir la suite ».
9) Interdire l'accès par l'Etat Général des Fonds à un inventaire dont il a été décidé que la publication est « Invisible mais accessible »

Amélioration de la recherche transversale par la base géographique dans les inventaires XML/EAD

.

Recherche et affichage des résultats en l'état actuel

Le chercheur accède à ces informations à partir d'un formulaire spécifique, différent de la recherche sur champ « Lieu » du formulaire de recherche avancée commun à l'ensemble des inventaires publiés, utilisant des liens webdynamiques (terme rattaché par une relation hiérarchique ou d'équivalence) ou par une recherche plain-texte simple.
Sa recherche porte aussi bien sur les termes d'indexation déjà existant que sur les termes présents dans ces informations et qui ne sont pas indexés dans un instrument de recherche. Ainsi le chercheur peut demander « Ho Chi Minh Ville » (jamais indexé en tant que tel) et accéder aux informations concernant « Saigon (Vietnam) », lui expliquant les raisons du changement de nom et la date à laquelle il est intervenu.

Il peut, à partir de cet écran, effectuer, au choix, les opérations suivantes :

1) sélectionner le terme d'indexation correspondant à sa requête, et lancer la recherche à partir de ce seul terme
2) en cas d'équivalences existantes, il peut :
    a. soit associer tous les termes d'indexation équivalents
    b. soit sélectionner un ou plusieurs termes d'indexation (par exemple choisir uniquement « Fort-Royal (Martinique, Île de la) » et non étendre aussi sa recherche à « Fort-de-France (Martinique; Île de la) », nouveau nom après la Révolution
3) sélectionner tout ou partie des niveaux immédiatement inférieurs s'ils existent, la recherche correspond soit à la totalité des termes d'indexation immédiatement inféodés, soit aux seuls termes choisis dans ces niveaux
4) déplier le (ou les) niveau(x) éventuels(s) supérieur(s) et sélectionner également tout ou partie des niveaux ainsi dépliés, la recherche correspond alors aux termes d'indexation ainsi choisis(s'ils existent)
5) revenir au formulaire de recherche initial

La recherche ainsi précisée (sur un ou plusieurs termes d'indexation) ramène toutes les occurrences y répondant. Les résultats, comme c'est déjà le cas pour toute recherche

transversale à l'ensemble des instruments de recherche, s'affichent par inventaire, avec indication du nombre d'occurrences répondant à la requête pour chacun d'eux.

Modifications attendues

1) Le chercheur devra pouvoir interroger cette base de connaissance non seulement à partir de l'expression exacte de la vedette indexée, mais également à partir d'une expression incomplète des termes d'indexation. Par exemple : demander « Nouvelle-Orléans » et accéder aux informations sur ce qui correspond au terme indexé : « La Nouvelle-Orléans (Louisiane ; États-Unis) ».
2) Ne pas reprendre en tête des variantes la vedette géographique titre de la notice (ou ne pas l'indiquer en variante lorsqu'il n'y a pas d'autre vedette associée). Actuellement le bouton du titre de fiche commande la sélection de toutes les vedettes et celui du titre répété en première variante permet sa sélection individuelle :
   Exemple :
   + **Fort-de France** (le signe « + » permet de sélectionner les quatre variantes pour le panier de lieux)
      + Fort-de-France (le signe « + » sélectionne seulement Fort-de-France pour le panier de lieux)
      + Fort-Royal (etc.)
      + Fort-de-la-République
      + République-Ville.

   Proposer un nouveau système de sélection.
3) Actuellement, toutes les vedettes géographiques titres de notice peuvent être sélectionnées pour le panier de lieux. Or toutes ne donnent pas accès à un inventaire en ligne puisqu'elles ont été créées pour instituer un niveau nécessaire dans l'organisation hiérarchique. Par exemple, ont été créés en « parents » les états des États-Unis, même si ils ne figurent pas en tant que tels dans l'instrument de recherche.
   Exemple : « Tennessee » a été créé pour pouvoir rattacher à cette entité géographique des noms de villes, cours d'eau, etc.
   Créer la possibilité de sélectionner uniquement les vedettes géographiques donnant accès à un inventaire.
4) Ajouter une possibilité de recherche en plein texte et proposer un écran d'affichage des résultats dans ce cas. En effet, il doit être possible, comme pour la base nominative « Personnes et familles », d'effectuer des recherches dans l'ensemble des informations saisies. Ainsi pour cette dernière base vous pouvez, dans une zone de saisie « plein texte » demander par exemple une recherche sur « architecte » et obtenir la liste de tous les architectes ou « photographe + Alger » et avoir la liste des photographes ayant exercé à Alger. Pour la base géographique ce pourrait être par exemple « seigneurie + Canada) » ou « habitation + Guadeloupe » et collecter ainsi toutes les seigneuries du Canada ou les habitations de la Guadeloupe y figurant.
5) Afficher, à côté des coordonnées GPS, une vue aérienne.
6) Afficher la date de création ou de dernière mise à jour de la fiche.
7) Créer la possibilité de sélectionner de façon simultanée tous les lieux-enfants
8) Créer un bouton d'aide. Les ANOM devront avoir la possibilité de rédiger et de modifier cette page.
9) Création dans l'outil de gestion de l'administrateur de la base de connaissances géographique d'une fonctionnalité permettant de savoir quelles sont les entrées dans la table géographique (outil propre aux ANOM) pour lesquelles il n'existe pas encore de fiches. En effet, il est souhaitable que l'administrateur dispose d'un contrôle permanent

de la cohérence entre les vedettes d'indexation et leur description dans la base de connaissance. Or il est toujours possible qu'une vedette ait été oubliée ou non signalée et qu'elle ne soit donc pas intégrée dans la base de connaissance. Un tel outil vise donc à écarter de telles possibilités.

Amélioration de la recherche transversale par vocabulaire aux inventaires XML/EAD

Le chercheur peut, à partir d'une liste déroulante, choisir un terme. La réponse est actuellement une page du dictionnaire (définition et domaines associés). Voir l'exemple page 7.

Les modifications à apporter sont les suivantes :

1) Offrir au chercheur la possibilité d'effectuer de 2 types de requête :
    a. Recherche de toutes les occurrences d'un terme (« café » dans l'exemple) quels que soient les domaines associés. Les résultats doivent être affichés par inventaire avec indication du nombre d'occurrences pour chacun d'eux.
    b. Recherche sur un terme (« café » dans l'exemple) associé à 1 ou plusieurs domaines dans un panier (« Végétaux », « Agriculture » etc. dans l'exemple). Les résultats doivent être affichés par inventaire avec indication du nombre d'occurrences pour chacun d'eux.
2) Mettre un bouton « Effacer ».
3) Offrir au chercheur la possibilité d'effectuer un recherche sur un ou plusieurs domaines. La liste alphabétique des domaines doit être proposée au chercheur au moyen d'une liste déroulante.
    a. S'il porte son choix sur un domaine, la liste des termes indexés appartenant à ce domaine doit lui apparaître. Il devra alors pouvoir exprimer sa recherche soit en souhaitant qu'elle porte sur l'ensemble de tous les termes, soit en sélectionnant ceux qui sont pertinents pour lui en les plaçant dans un panier. Les résultats doivent être affichés par inventaire avec indication du nombre d'occurrences pour chacun d'eux.
    b. Il doit également être possible de croiser plusieurs domaines (« Architecture religieuse et ornements religieux » et « Cultes et religions », par exemple). Les résultats doivent être affichés par inventaire avec indication du nombre d'occurrences pour chacun d'eux.

Modification de l'outil attribuant les identifiants ARK

Lors de la rédaction d'un instrument de recherche, il est fréquent que son auteur effectue des renvois à d'autres parties de cet inventaire. Ceux-ci, conformément à la DTD EAD, se distinguent selon que les renvois aient lieu à l'intérieur d'une même instance (fichier) EAD, ou d'une instance à une autre.

Dans le premier cas l'élément « ref » est utilisé et un attribut TARGET, alimenté par la valeur de l'attribut ID du composant visé permet d'atteindre celui-ci.

Exemple :

C1 = composant cible = <c level="item" id="cc7A63f19">

C2 = composant source = Envoi d'un <ref target="cc7A63f19">mémoire</ref>

Dans le second cas l'élément <archref> est un attribut HREF accueil la référence de l'instance EAD cible complétée par l'attribut ID permettant d'aller directement au composant visé.

Exemple :

C1 = composant cible = <c level="item" id="cc7A61f68">

C2 = composant source = <archref href="Lacrosse.xml#cc7A61f68">emprunt fait par le général Lacrosse au gouverneur</archref>

L'auteur est donc contraint de forger des identifiants qui ne sont pas des identifiants pérennes de type ARK pour pouvoir effectuer les renvois nécessaires. Une fois l'inventaire achevé et remis à l'administrateur des inventaires encodés en XML/EAD, ce dernier incrémente les identifiants ARK automatiquement, mais il est obligé de traiter manuellement tous les identifiants provisoires insérés par l'auteur pour les convertir en identifiants ARK.

Il conviendrait de modifier l'outil attribuant les identifiants ARK dans les fichiers XML/EAD pour qu'il puisse traiter automatiquement le remplacement des identifiants provisoires de façon à éviter une intervention manuelle assez longue et délicate (pour certains inventaires volumineux les renvois peuvent atteindre plusieurs centaines).

Intégration d'un outil de gestion et de publication des notices d'autorités encodées en XML/EAC-CPF dans IREL.

Les Archives nationales d'outre-mer souhaitent développer la production de notices d'autorités encodées en XML selon le schéma EAC-CPF (version 2010). Une première expérience de production de notices portant sur les acteurs importants de la colonisation a déjà été menée aux ANOM, qui, pour l'occasion, avaient développé en interne une application permettant la rédaction de notices XML selon la DTD EAC (version 2004).
Ces notices, qui portent essentiellement sur des gouverneurs d'Ancien Régime, sont accessibles via la base nominative « Personnes et familles ». L'abandon de la DTD au profit du schéma EAC-CPF a eu pour conséquence l'arrêt de leur production. Cependant, le développement récent par le Conseil International des Archives d'un logiciel de production (ATOM) selon le schéma EAC-CPF et pour lequel les ANOM ont participé aux tests effectués par la France, va conduire le service à reprendre la production de ces notices d'autorité.

Entrepôt Subversion (SVN) pour le notices EAC-CPF

Les ANOM disposent d'un entrepôt spécifique pour les notices d'autorité. Il paraît nécessaire de séparer les notices Personnes et Familles et les notices Personnes morales en 2 entrepôts distincts. Il faudra également récupérer les données contenues dans les notices déjà élaborées suisvant la DTD pour les faire migrer vers le schéma EAC.

Création d'un outil de publication des notices EAC-CPF

Comme pour les inventaires encodés en XML/EAD, les ANOM devront disposer d'un outil de publication permettant de contrôler les notices sur un serveur test (servirel) et de les rendre publiques après validation.

Adaptation de l'affichage actuel des notices en DTD EAC aux notices EAC-CPF

De nombreux changements sont intervenus, aussi bien sur la structure que sur les éléments et leurs attributs, entre la version DTD et la version schéma. Un même type d'affichage doit être proposé au chercheur. Il conviendrait d' appliquer aux notices rédigées en EAC-CPF (version 2010) les caractéristiques d'affichage des notices actuelles (EAC version 2004) et des fonctionnalités qui y sont liées.

Intégration dans la page d'accueil d'IREL d'outils d'accès aux notices d'autorité

Les ANOM souhaitent disposer d'un outil leur permettant d'inclure, au fur et à mesure de leur élaboration, les domaines touchés par ces notices dans la page d'accueil d'IREL, au même niveau que les outils de recherches transversales (base nominative « Personnes et familles », « Vocabulaire », « Base de connaissance géographique ».) Dans un premier temps, seuls seront concernés les compagnies de commerce et les ordres religieux puisque les patronymes ont déjà leur propre outil de recherche.
La recherche sur les notices d'autorité à l'intérieur d'un domaine (par exemple les ordres religieux) doit pouvoir s'effectuer à partir de la liste de la forme normalisée des vedettes. Pour les ordres religieux par exemple ce pourrait être :

- Augustines
- Augustins
- Bénédictins
- Capucins
- Etc.

La notice affichée donne accès à la liste des inventaires dans lesquels apparaît un ordre religieux, et pour chaque inventaire au nombre d'occurrences.

Amélioration de la recherche transversale par la base nominative « Personnes et familles »

1) Créer un panier permettant au chercheur une recherche sur plusieurs noms.
2) Transformer les renvois informatifs actuels en liens. La base nominative Personnes et Familles http://anom.archivesnationales.culture.gouv.fr/nominatif/ est une base de données PHP/MySQL qui liste toutes les personnes citées dans les instruments de recherche XML/EAD. Les données comportent, entre autres, une colonne [Nom] et une colonne [Prénoms].
   Afin de faciliter la recherche des renvois vers d'autres formes du nom complètent les données. Ces renvois sont de deux types : « voir » et « voir aussi ». « voir » renvoie vers un patronyme précis (exemple La Fayette, voir Du Motier, Marie Joseph Paul Yves Roch Gilbert). « voir aussi » renvoie vers d'autres formes d'écriture d'un nom qui peuvent concerner différentes personnes du même patronyme (exemple Entremont, voir aussi Antremont). Les renvois sont alors placés dans la colonne [prénoms].
   Amélioration de la navigation : il est demandé de transformer ces renvois, jusqu'alors informatifs, en liens renvoyant vers les enregistrements concernés.
3) Surligner le résultat de la recherche dans l'IR quand la recherche se fait à partir de la base nominative.

Amélioration de l'application de l'Etat civil

Les sources concernant l'état civil font l'objet d'une application autonome qui permet, outre la recherche, la visualisation en ligne des actes eux-mêmes.

http://anom.archivesnationales.culture.gouv.fr/caomec2/

Situation actuelle

Il s'agit d'une base de données en PHP/MySQL.

Structure du fichier

Le fichier CSV de données les six champs suivants :
- Territoire (nom de l'ancienne colonie)
- Commune
- Année
- Type d'acte
- Nature du registre
- Nom de l'image

Le type d'acte est encodé de la façon suivante :
- Naissance.....AC NA
- Reconnaissance.....AC RE
- Mariage.....AC MA
- Divorce.....AC DI
- Décès.....AC DE
- Enfant mort-né.....AC SV
- Jugement.....AC JU
- Affranchissement……AC AF

Un critère de choix, nommé « Thème » est possible pour certains territoires. Ce ne sont pas des actes d'état civil proprement dits, mais plutôt des informations spécifiques à certains territoires. Ils sont indexés de la façon suivante : TH ES (esclavage) en doublant l'enregistrement en modifiant le champ « Type d'acte ».
Les tables annuelles suivent les mêmes règles que les actes.

Inscription des actes dans les registres

Les actes sont inscrits dans les registres de deux manières : soit par catégorie, soit tous

mélangés, ce qui influe sur la manière d'alimenter le fichier.

Actes regroupés par type

Les reconnaissances et les affranchissements sont inscrits avec les naissances. L'enregistrement est alors doublé en modifiant le champ « Type d'acte » (un enregistrement avec AC NA, doublé par un enregistrement avec AC RE ou AC AF).

Les divorces sont inscrits avec les mariages. L'enregistrement est alors doublé en modifiant le champ « Type d'acte ».

Les enfants mort-nés sont inscrits avec les décès. L'enregistrement est alors doublé en modifiant le champ « Type d'acte »

S'il est fait mention dans un acte d'un esclave, l'enregistrement est alors doublé en modifiant le champ « Type d'acte » par TH ES.

Actes mélangés

Les vues sont alors indexées dans le champ « Type d'acte » par AC TS.

Les enregistrements sont alors doublés en modifiant le champ « Type d'acte » pour indiquer les actes spécifiques (reconnaissance, affranchissement, divorce, jugement, enfants mort-né, esclavage).

Formulaire d'interrogation de la base

Plusieurs critères de choix sont proposés :

- Territoire. Le choix du territoire peut se faire par l'intermédiaire de la planisphère ou par le formulaire d'interrogation qui offre une liste déroulante. Il est obligatoire.
- Commune : elle est choisie dans une liste déroulante comprenant toutes les communes du territoire concerné.
- Type d'acte : il est choisi dans une liste déroulante.
- Date : ce peut être une année précise ou un intervalle entre deux années (année début et année fin).

Un tri est proposé : par commune, par nom, par année.

Modifications attendues

1. Affichage des résultats

1. L'indexation tous actes « AC_TS » comporte les types d'acte suivants AC_NA, AC_MA , AC_De

Lorsque que la demande porte sur un de ces types d'actes (AC_NA, AC_MA , AC_DE) et que la réponse se trouve dans un fichier indexé AC_TS (tous actes) le résultat qui apparaît dans le tableau de réponses est la mention « Naissance »à la requête naissance, « Mariage » à la requête mariage et « Décès » à la requête décès.
Pour répondre plus justement au chercheur, il faut faire apparaître en réponse « Tous actes ».

2. Pour l'Algérie, la recherche porte d'abord sur l'identifiant AC_NA même si les

actes spécifiques comme les reconnaissances sont bien indexées dans la table des types d'actes AC_RE. Par conséquent quand la requête porte par exemple sur les reconnaissances, il est indiqué au chercheur qu'il n'y a aucune réponse à la demande. Il faut donc que la recherche porte sur l'identifiant AC_NA puis sur le type d'acte AC_RE.

2. Validation des fichiers de chargement

Avant import définitif dans la base de données les fichiers d'alimentation sont chargés dans une table temporaire munie d'un filtre spécifique d'importation qui permet, entre autres, de s'assurer de la présence d'un fichier image pour chaque enregistrement du fichier d'alimentation.

Pour le territoire Algérie ce filtre permet de vérifier ou de détecter :

- la présence ou l'absence d'image sur le serveur

- un mauvais séparateur de ligne (par défaut CR LF, ou \r\n)

- un mauvais séparateur de colonnes (par défaut le point virgule)

- l'ordre des colonnes ou de cellules vides

- les types d'actes inconnus (non déclarés dans la table _typeacte)

- un territoire inconnu (non déclaré dans la table _territoire)

Pour les autres territoires certaines erreurs de validation lors de l'importation semblent avoir été omises notamment la validation des types d'actes.

Il est donc demandé de corriger le filtre d'importation afin de détecter les erreurs de catégories d'actes permettant ainsi de s'assurer de la cohérence des requêtes.